Manuscript No: PLPSAMIP.TEX

Atmospheric Global Climate Models' Representation of the Landsurface: From AMIP I to AMIP II

A. Henderson-Sellers

Director, Environment

Australian Nuclear Science and Technology Organisation

Lucas Heights, NSW, Australia

12 December 1999

Address for Correspondence

Professor A. Henderson-Sellers
Director, Environment
Australian Nuclear Science and Technology Organisation
Lucas Heights Science and Technology Centre
Private Mail Bag 1
Menai
NSW 2234
AUSTRALIA
email: ahs@ansto.gov.au
fax: +61 2 9717 3599

Abstract

Intercomparison of the present-day simulations of between 10 and 20 landsurface parameterization schemes as part of AMIP I and in associated offline experiments (with prescribed meteorology) identified strengths and weaknesses in landsurface representation in global atmospheric models. Specifically, AMIP I diagnostic subproject number 12 showed that: (i) every landsurface simulation was an outlier in some characterization: no one "best" model/scheme combination emerged (Love and Henderson-Sellers, 1994); (ii) some simulations displayed obvious flaws including failure to conserve continental moisture/energy and trends in moisture stores due to inadequate model initialization procedures and to coding errors (Love *et al.*, 1995); and (iii) the scatter in regional-scale energy/moisture partitionings among the coupled simulations was substantially greater than in comparable off-line experiments, in spite of the presence of two-way feedback enabled by AMIP (Irannejad *et al.*, 1995; Qu and Henderson-Sellers, 1998). These results, although agreed within the community, have not been welcomed. The coding and coupling errors were, naturally, not widely publicized although privately some thanks were received. The inadequacies of experimental design have been recognized and corrected in AMIP II. The findings of great diversity in the characterization of the continental surface climate has, to a large extent, been dismissed by the atmospheric modelling community as being the result of (a) known differences among landsurface schemes or (b) intercomparison of off-line and coupled simulations. Here, we highlight some findings from AMIP I and ponder the likely fate of outcomes from AMIP II.

1. Landsurface Processes in AMIP I

The first phase of the Atmospheric Model Intercomparison Project (AMIP I) drew into a central archive simulations of the period 1979–1988 made by about 30 Atmospheric General Circulation Models (AGCMs). One of the many diagnostic subprojects (number 12) was led by the Project for Intercomparison of Landsurface Parameterization Schemes (PILPS) (Gates, 1992; Love and Henderson-Sellers, 1994). PILPS is a World Climate Research Programme (WCRP) project operating under the auspices of the Global Energy and Water Cycle Experiment (GEWEX) and the Working Group on Numerical Experimentation (WGNE) (Henderson-Sellers and Brown, 1992; Henderson-Sellers *et al.*, 1993, 1996; Pitman *et al.*, 1993, 1999). Its goal is to enhance understanding of the parameterization of fluxes of heat, moisture and momentum between the atmosphere and the continental surface in climate and weather forecast models. From 1990 to 1996, diagnostic subproject 12 contributed analyses of AMIP I results of relevance to the characterization of the continental climate.

In the AMIP I global models, the landsurface schemes were solely boundaries for energy and water exchanges between the air and the land. Most landsurface shcemes were well established, dating back to Manabe (1969) and were generally believed to be working well. The AMIP I experiment was designed to include evaluations of the landsurface. Despite this, there were a number of difficulties associated with the analysis. These included:

1. the failure to initialize soil moisture consistently among AMIP I models (this was simply forgotten);

2. a rather restricted set of "standard output" variables (this was a known drawback but deemed acceptable);
3. the limited variety of landsurface schemes (LSS) in the AGCMs of that time (both a strength and a weakness); and
4. poor quality control of archived results and relatively poor documentation of the LSSs which were represented (this occurred despite very significant efforts at PCMDI and PILPS 'Central').

Despite these obstacles and inhibiting factors, it was possible to undertake evaluations of the simulated landsurface climates. Although it was originally intended to concentrate on "validation", the outcomes of diagnostic subproject 12 of AMIP I are perhaps best described as "learning on the job". They are described in detail in PCMDI report No. 12 (Love and Henderson-Sellers, 1994) and in a variety of papers (e.g. Henderson-Sellers *et al.*, 1996; Qu and Henderson-Sellers, 1998).

The AMIP I/12 results can be summarized into three main findings:

1) No "best" landsurface simulation could be identified: every surface climate was an outlier in some respect (Love and Henderson-Sellers, 1994);
2) results revealed serious errors of execution such as nonconservation of continental moisture and/or energy and pronounced trends in moisture stores which were traced back to coding/coupling mistakes and incorrect/inadequate initialization (Love *et al.*, 1995); and
3) at a regional scale, energy and moisture partitions among the coupled simulations was substantially greater than in comparable off-line experiments, suggesting that the hypothesis that two-way feedbacks between land and atmosphere dampen landsurface

climate differences is incorect or at least unproven (Irannejad *et al.*, 1995; Qu and Henderson-Sellers, 1998).

The twofold outcome of AMIP I/12 was: (i) landsurface codes were not working especially well and certainly not as well as the AGCM owners and users believed and (ii) the neglect of stomatal resistance parameterization in incorporated LSSs rendered the representation of many continental climates rather poor. Results suggest that at least two years' simulation, and often much longer, is required to overcome the difficulties of soil moisture initialization. Energy and moisture differences over multi-year periods were caused in AMIP I by the philosophies underpinning the incorporation of deep soil processes, as a result of poor experimental design, and by coding errors.

2. Energy and Moisture Inconsistencies at the Landsurface

In PILPS Phase 1, participating landsurface schemes were integrated for many years using synthetic meteorological forcing supplied at 30 minute timesteps from a global climate model. A single year's meteorology was used for as many annual cycles as was required for the particular landsurface scheme to come into equilibrium with the prescribed atmospheric conditions. In these experiments, there is no feedback to the atmosphere. Results from the first series of these offline intercomparisons (Phase 1(a)) showed disturbingly large differences (Pitman *et al.*, 1993, 1999).

Figure 1 shows the ranges of PILPS Phase 1(a) and 1(b) results (cross bars) superimposed upon the 1(b) scatter plots (crosses) of annually-averaged evapotranspiration against annually-averaged sensible heat flux for two of the situations evaluated: a tropical forest and a mid-latitude grassland. Phase 1(b) is a repeat of 1(a) using more stringent prescriptions and more careful reporting to try to remove the scatter in the 1(a) results. The Phase 1 results are from equilibrium surface climates (i.e. after many years of integration), have the same descriptions of the surface vegetation and soil and identical forcing meteorology. Therefore, the net radiant energy at the surface should be similar (differing only as a result of different calculations of surface temperature) and differences should be manifested primarily as different partitions between sensible and latent heat fluxes i.e. the crosses should lie close to a diagonal line. The scatter (away from the diagonal) in Figure 1 (around 10 W m^{-2}) must represent either differences in the net radiant energy (probably differences in the specification of surface albedo and computed temperature) or errors in the surface energy budget. The range (length of the diagonal) between participating schemes shows how different representations of the continental surface differently partition net absorbed energy.

The range in latent heat fluxes is about 100 W m^{-2} for tropical forest and about 50 W m^{-2} for the grassland with commensurate ranges in the sensible heat fluxes. Monthly and diurnal ranges are larger (e.g. Pitman *et al.*, 1993, 1999) and are found in all diagnosed characteristics of the surface climate. Phase 1 allowed only intercomparisons, there are no "correct answers", but results show that no model lies close to the central values in all experiments/variables. The simpler (bucket) schemes tend to evaporate more than schemes which include a representation of canopy processes or at least a parameterization of stomatal resistance, and lie towards the lower right hand of Figure 1.

The PILPS Phase 1 intercomparisons also showed that most landsurface schemes required many years to come to thermal and hydrologic equilibrium with the forcing meteorology and that the time needed and the final equilibrated state could differ depending upon the initialization of the moisture stores: spin-up from completely arid conditions generated a different equilibrium surface climate from a dry-down from completely saturated conditions. The equilibration time is a function of the maximum available soil moisture except in schemes where the lowest layer is saturated. All models required at least two years to equilibrate, many required greatly in excess of 10 years and some hundreds of years (Yang *et al.*, 1995). Thus short period (less than 10 years) simulations of landsurface disturbances (e.g. tropical deforestation and desertification) will produce a range of results dependent upon initialization and the surface scheme. More recent PILPS' experiments (in Phase 2) use observations from real sites. For example, Cabauw, the Netherlands, was selected because it has saturated deep soil year-round, so overcoming soil moisture initialization differences (Chen *et al.*, 1997). The range amongst the predictions as compared with observed turbulent energy fluxes and the scatter about the zero total energy budget line continue to be a concern as do arbitrary specification of soil depths and hence soil moisture stores (Slater *et al.*, 2000).

For AMIP I/12, it was hypothesized that the range amongst the landsurface schemes' predictions would be greatly decreased once they were incorporated into atmospheric models so that feedbacks existed. It has proved possible to evaluate a subset of landsurface schemes in this coupled mode as part of the AMIP I diagnostic project number 12. In this case, atmospheric general circulation models were integrated for the 10 years from 1979 to 1988 using prescribed sea surface temperatures and sea-ice distributions (Gates, 1992). Although the simulations were supposed to be as similar to one another as possible (e.g. the same solar flux and trace gas amounts), the need to initialize soil moisture consistently had been overlooked. Many AMIP models show a one to two year spin-up period in their surface energy and moisture budgets (Love and Henderson-Sellers, 1994).

The AMIP models' climates differ from the synthetic conditions used in the PILPS Phase 1 offline experiments. In particular, total precipitation amounts in both locations are only 50–60% of the offline prescription and incident radiation totals differ. The different "forcing" accounts for some of the differences between the AMIP I/12 and PILPS Phase 1 results e.g. the decrease in evaporation by about 20 W m^{-2} in the tropical forest (Figure 1(a)). Differences are also due to feedbacks e.g. for grassland, the sensible heat flux becomes positive (i.e. energy is lost to the atmosphere in Phase 3) while the evaporative flux changes little (Figure 1(b)). While these differences are interesting, Figure 1 demonstrates that the range in annually-averaged sensible and latent heat fluxes in AMIP I/12 are as great as in the offline experiments. Feedbacks between the atmosphere and the surface do not reduce the large range in predictions of landsurface climates.

AMIP I/12 intercomparisons have permitted the calculation of continental surface energy residuals: the sum of the net shortwave, net longwave, sensible and latent heat fluxes. Unless there are other significant energy terms e.g. heat storage or loss in the soil, this energy residual should be zero when in equilibrium (PILPS off-line) or when averaged over a number of years (AMIP I). Many LSSs have non-zero surface energy residuals: hemispheric mean values range from 2.8 to 4.4 W m^{-2} and maxima exceed 10 W m^{-2} (Table 1). Whilst residuals up to 1 or perhaps 2 W m^{-2} might be attributable to different values employed for latent heats or lack of complete equilibrium, the large values found here are disturbing.

This energy residual is observed or can be seen in published work (Garratt *et al.*, 1993) but is rarely discussed except for the suggestion that it is due to incomplete calculations in the processes of snow melt or longer equilibration times in high latitudes (Yang *et al.*, 1995). The AMIP I/12 results do not support these hypotheses. Although the northern hemisphere mean and range are somewhat greater than those for the southern hemisphere (Table 1), point values, with and without snow, are similar and there is no consistent spatial pattern of larger values at high latitudes (Figure 2). AMIP I/12 energy residuals show little inter-annual behaviour but have a very strong annual cycle: positive in summer and (smaller) negative in winter. The amplitude of this cycle is larger for large values of the energy residual. The energy residual is larger for more complex landsurface schemes, although the number of models considered here is too small to draw any firm conclusions. At least one residual was due to a coding error which caused the evaporation seen by the landsurface scheme to be underestimated by about 7%.

Other energy residuals may be artifacts of the landsurface parameterizations. Most landsurface schemes employ a zero flux formulation, no heat transfer at the bottom of the deepest soil layer (Deardorff, 1978), or fix the temperature of their deepest soil layer. If the specified temperature is not equal to the long-term average of the scheme's surface temperature, then there will be a net energy flux into or out of the soil[1]. This may explain one other large energy residual in AMIP I/12. The others remain unexplained.

[1] For a heat conductivity of 1 W m^{-1} K^{-1} and two bottom layers of thickness 0.5 m, an error of 1 K will lead to a transfer of energy of 2 W m^{-2}. Oke (1978) gives values of heat conductivity in the range 0.06 to 2.2 W m^{-1} K^{-1} for soil types peat, sandy and clay, including dry and saturated (Oke, 1978).

3. Do Landsurface Climate Simulations Matter?

The AMIP/PILPS (AMIP I/12) experiments found disturbingly large differences among results from participating landsurface schemes and significant discrepancies from anticipated budgets. In contradiction to the hypothesis that off-line experiments exaggerate differences due to different LSSs, the range in AMIP I/12 is not diminished despite coupling AGCMs and LSSs. Some of the range and scatter is due to coding errors. Other differences are due to differences in the philosophies underlying the LSSs and the coupling (AGCM to LSS) codes. The scatter in annually-averaged surface fluxes is commensurate with commonly imposed enhanced greenhouse forcing and the range among schemes is about an order of magnitude larger; monthly means and diurnal samples differ even more (e.g. Gates *et al.*, 1996 cf. Henderson-Sellers *et al.*, 1996).

A decade ago, when AMIP I was conceived, these results would have been view as of little or no interest. First generation LSSs (e.g. Manabe, 1969) were very simple: the surface hydrology was a 'bucket' from which evaporation occurred 'on demand' and 'run-off' quite literally disappeared. These were the predominant schemes in AMIP I. Second generation LSSs were developed during the 1980s (e.g. Sellers *et al.*, 1986; Dickinson *et al.*, 1986) to incorporate more physically realistic interactions of energy and water at the land surface. Some aspects of these schemes have been shown to be of importance for weather forecasting and climate simulation (e.g. Viterbo and Beljaars, 1995; Betts *et al.*, 1998; Sellers *et al.*, 1997). One or two such schemes were present in AMIP I and many more have been represented in other phases of PILPS. It is likely that AMIP II will include larger numbers of such schemes and hence a greater variety of landsurface characterization.

Third generation LSSs combine the physical processes of energy and water exchange with the biophysical exchanges needed to represent photosynthesis, respiration and, in some schemes, decay (e.g. Xiao *et al.*, 1998; Tian *et al.*, 1999). While these biophysical processes are known to be essential for climate modelling, their addition is not needed or sought for weather forecasting. As AMIP II may also include a few third generation LSSs, the effect of the inclusion of biophysical processes on atmospheric simulations may be to spread the characterization of landsurface climates still further. It also suggests that the experimental design of AMIP II may suffer from, as yet unknown, initialization problems.

Unfortunately, it seems that the range of types of landsurface schemes is likely to be as much of a difficulty in AMIP II as in AMIP I. Similarly, there is no obvious or documented reason to believe that coding and coupling errors of the type that gave rise to the features of Table 1 and Figure 2 will not be present in the AMIP II. Emphasis for many atmospheric modellers has been, and for some still is, on the weather simulated by their models. The 1993 Mississippi floods, forecast by the ECMWF medium range model, were found to be sensitive to the landsurface parameterization scheme employed (Betts, 1994) focussing the attention of weather forecasters on more appropriate means of representing energy and moisture components of the continental surface. Moreover, *climate* scientists cannot and do not wish to neglect the landsurface for a variety of additional reasons including:

i) as climate simulations are now for hundreds, and sometimes thousands, of years, the assumption (or pretence) that the continents do not change is untenable;

ii) alterations to the atmosphere (e.g. the addition of CO_2) are known to prompt direct responses in landsurface character (e.g. stomatal closure and more rapid plant growth if other conditions are favourable);

iii) as carbon budgetting is a critical component of modern climate science landsurfaces can no longer be treated as energy, water and momentum exchanges only; and

iv) other subcomponents of modern climate models demand inputs from landsurfaces e.g. runoff delivering fresh water to ocean models; atmospheric gases emanating from or being absorbed by vegetation and soils.

While most landsurface schemes now incorporate at least a simple representation of stomatal resistance to evaporation as a function of soil moisture and routing of runoff, the focus of atmospheric modellers on weather and circulation patterns may mean that AMIP II does include many AGCMs coupled to biospherically-realistic LSSs. Even if some are represented their presence may hinder analysis and inhibit interpretation and/or acceptance of results in atmospheric circles.

Many of the "surprises" currently being reviewed by the Intergovernmental Panel on Climate Change (IPCC) for inclusion in their Third Assessment Report draw on biophysical interactions between the landsurface scheme simulations and those of the other components of modern climate models. Suggestions for climate surprises include:

a) daytime stomatal closure globally could increase diurnal temperature ranges new the landsurface as a result of reduced daytime transpirational cooling;

b) stomatal closure in the tropics could augment direct greenhouse warming regionally;

c) boreal forests and wetlands which dry and/or warm could release extra carbon (as CO_2 or CH_4) to the atmosphere; and

d) landsurface feedbacks could enhance drought conditions in mid-latitude continents, prompting or encouraging 'climate lock' over multi-month periods.

While the likelihood of these “surprises” occurring cannot yet be assessed, most risk analyses are being conducted using fully coupled climate models incorporating adequately complex landsurface schemes.

The results from AMIP I/12 were dismissed by some as presenting too large a range of inter-scheme differences. The preferred explanation was that PILPS had incorporated LSSs of a wider range of complexity than AMIP I and therefore this had contributed to the diversity of results. This may be correct. If it is, it will not be confined to AMIP I. AMIP II, the Coupled Model Intercomparison Project (CMIP) and the Past Model Intercomparison Project (PMIP) include GCMs which have (optionally) LSS(s) that capture many aspects of a fully interactive vegetative biosphere on land i.e. third generation biophysical schemes. The extent to which aspects of carbon uptake and release; vegetation change and even ecosystem evolution are incorporated in current LSSs is likely to give rise to at least as large a population diversity in AMIP II as was found in AMIP I.

The challenge for AMIP II may be how it can best serve the landsurface and climate modelling communities. Despite small efforts with one or two models (e.g. Gedney *et al.*, 1999) and in CMIP, AMIP remains the most comprehensive intercomparison project for global models. The AMIP II models ought to include more appropriate LSSs than the Manabe (1969) bucket. There is therefore a possibility that some of the topics of interest to climate modellers (e.g. (d) above) can be addressed by analysis of AMIP II results. On the other hand, the continuing preference among some modelling groups for atmospheric validation above all other components may hinder the total value deliverable by AMIP II to e.g. the IPCC with respect to landsurface climate simulation.

Acknowledgements

This report draws on the results of the AMIP I Diagnostic Subproject Number 12 and other PILPS experiments. All the contributors to these efforts are gratefully acknowledged.

References

Betts, A.K., 1994, Presentation at AMS Conference, Global Change Symposium, Nashville, Tennessee

Betts, A.K., Viterbo, P. and Wood, E., 1998, Surface energy and water balance for the Arkansas–Red River basin from the ECMWF reanalysis, *J. Climate,* **11**, 2881–2897

Chen, T.H., Henderson-Sellers, A., Milly, P.C.D. Pitman, A.J., Beljaars, A.C.M., Polcher, J., Abramopoulos, F., Boone, A., Chang, S., Chen, F., Dai, Y., Desborough, C.A., Dickinson, R.E., Dumenil, L., Ek, M.B., Garratt, J.R., Gedney, N., Gusev, Y.M., King, J., Koster, R., Kowalczyk, E., Laval, K., Lean, J., Lettenmaier, D., Liang, X., Mahfouf, J-F., Megelkamp, H.-T., Mitchell, K., Nasonova, O.N., Noilhan, J., Robock, A., Rosenzweig, C., Schaake, J., Schlosser, A., Schulz, J.P., Shao, Y., Shmaking, A.B., Verseghy, D.L., Wetzel, P., Wood, E.F., Xue, Y., Yang, Z-L. and Zeng, Q., 1997, Cabauw experimental results from the Project for Intercomparison of Landsurface Parameterization Schemes, *J. Climate,* **10(7)**, 1194–1215

Deardorff, J.W., 1978, Efficient prediction of ground surface temperature and moisture with inclusion of a layer of vegetation, *J. Geophys. Res.,* **83(C4)**, 1889–1903

Dickinson, R.E., Henderson-Sellers, A., Kennedy, P.J. and Wilson, M.F., 1986, Biosphere-atmosphere transfer scheme (BATS) for the NCAR community climate model, *NCAR Technical Note,* **NCAR TN–275+STR**, 69pp

Garratt, J.R., Krummel, P.B., Kowalczyk, E.A., 1993, The surface energy balance at local and regional scales — a comparison of general circulation model results with observations, *J. Clim.,* **6**, 1090–1109

Gates, W.L., 1992, The atmospheric model intercomparison project, *Bull. Amer. Meteor. Soc.,* **73**, 1962–1970

Gates, W.L., Henderson-Sellers, A., Boer, G.J., Folland, C.K., Kitoh, A., McAvaney, B.J., Semazzi, F., Smith, N., Weaver, A.J. and Zeng, Q.-C., 1996, Climate models — evaluation, Chapter 5 **in** *Climate Change 1995. The Science of Climate Change* (eds. J.T. Houghton, L.G. Meira Filho, B.A. Callander, N. Harris, A. Kattenberg and K. Maskell), Cambridge University Press, Cambridge, 229–284

Gedney, N., Cox, P.M., Douville, H., Polcher, J. and Valdes, P.J., 1999, Characterising GCM land surface schemes to understand their responses to climate change, Hadley Centre Technical Note HCTN 4, United Kingdom Meteorological Office, Bracknell, Berks, RG12 2SY, UK.

Henderson-Sellers, A. and Brown, V.B., 1992, PILPS: Project for Intercomparison of Land-surface Parameterization Schemes. Workshop Report and First Science Plan, IGPO Publication Series No. 5, Science and Technology Corporation, Hampton, VA, 51pp

Henderson-Sellers, A., Yang, Z.-L. and Dickinson, R.E., 1993, The Project for Intercomparison of Land-surface Parameterization Schemes (PILPS), *Bull. Amer. Meteor. Soc.,* **74(7)**, 1335–1349

Henderson-Sellers, A., McGuffie, K. and Pitman, A.J., 1996, The Project for Intercomparison of Land-surface Parametrization Schemes (PILPS): 1992 to 1995, *Climate Dynamics,* **12**, 849–859

Irannejad, P., Henderson-Sellers, A., Shao, Y. and Love, P.K., 1995, Comparison of AMIP and PILPS off-line landsurface simulations, *Procs. of The First International AMIP Scientific Conference* (Monterey, CA, USA, 15–19 May 1995) (ed. W.L. Gates), WMO/TD–No. 732, WCRP, Geneva, 465–470

Love, P.K. and Henderson-Sellers, A., 1994, Land surface climatologies of AMIP–PILPS models and identification of regions for investigation, PCMDI Report 12, Lawrence Livermore Laboratory, California, 83pp

Love, P.K., Henderson-Sellers, A. and Irannejad, P., 1995, AMIP diagnostic subproject 12 (PILPS Phase 3): land-surface processes, *Procs. of The First International AMIP Scientific Conference* (Monterey, CA, USA, 15–19 May 1995) (ed. W.L. Gates), WMO/TD–No. 732, WCRP, Geneva, 101–106

Manabe, S., 1969, Climate and the ocean circulation: I, the atmospheric circulation and the hydrology of the earth's surface, *Mon. Wea. Rev.,* **97**, 739–774

Oke, T.R., 1978, *Boundary Layer Climates*, Methuen, 372pp

Pitman, A.J., Henderson-Sellers, A., Abramopoulos, F., Avissar, R., Bonan, G., Boone, A., Dickinson, R.E., Ek, M., Entekhabi, D., Famiglietti, J., Garratt, J.R., Frech, M., Hahmann, A., Koster, R., Kowalczyk, E., Laval, K., Lean, J., Lee, T.J., Lettenmaier, D., Liang, X., Mahfouf, J-F., Mahrt, L., Milly, P.C.D., Mitchell, K., de Noblet, N., Noilhan, J., Pan, H., Pielke, R., Robock, A., Rosenzweig, C., Schlosser, C.A., Scott, R., Suarez, M., Thompson, S., Verseghy, D., Wetzel, P., Wood, E., Xue, Y., Yang, Z-L and Zhang, L., 1993, *Project for Intercomparison of Land-Surface Parameterization Schemes (PILPS). Results from off-line control simulations (Phase 1A)*, GEWEX Report. IGPO Publication Series No. 7, 47pp

Pitman, A.J., Henderson-Sellers, A., Desborough, C.E., Yang, Z.-L., Abramopoulos, F., Boone, A., Dickinson, R.E., Gedney, N., Koster, R., Kowalczyk, E., Lettenmaier, D., Liang, X., Mahfouf, J.-F., Noilhan, J., Polcher, J., Qu, W., Robock, A., Rosenzweig, C., Schlosser, C.A., Shmakin, A.B., Smith, J., Suarez, M., Verseghy, D., Wetzel, P., Wood, E. and Xue, Y., 1999, Key results and implications from phase 1(c) of the Project for Intercomparison of Land-surface parameterization schemes, *Climate Dynamics,* **15**, 673–684

Qu, W. and Henderson-Sellers, A., 1998, Comparing the scatter in PILPS off-line experiments with that in AMIP I coupled experiments, *Global and Planetary Change*, **19(1–4)**, 209–223

Sellers, P.J., Mintz, Y., Sud, Y.C. and A. Dulcher, 1986, A simple biosphere model (SiB) for use within general circulation models, *J. Atmos. Sci.,* **43**, 505–531.

Sellers, P.J., Dickinson, R.E., Randall, D.A., Betts, A.K., Hall, F.G., Berry, J.A., Collatz, G.J., Denning, A.S., Mooney, H.A., Nobre, C.A., Sato, N., Field, C.B. and Henderson-Sellers, A., 1997, Modeling the exchange of energy, water and carbon between continents and the atmosphere, *Science,* **275**, 502–509

Slater, A.G., Schlosser, C.A., Desborough, C.E., Pitman, A.J., Henderson-Sellers, A., Robock, A., Vinnikov, K.Ya., Speranskaya, N.A., Mitchell, K., Boone, A., Braden, H., Chen, F., Cox, P., de Rosnay, P., Dickinson, R.E., Dai, Y.-J., Duan, Q., Entin, J., Etchevers, P., Gedney, N., Gusev, Ye.M., Habets, F., Kim, J., Koren, V., Kowalczyk, E., Nasonova, O.N., Noilhan, J., Schaake, J., Shmakin, A.B., Smirnova, T., Verseghy, D., Wetzel, P., Xue, Y., Yang, Z.-L and Zeng, Q., 2000, The representation of snow in land-surface schemes; results from PILPS 2(d), submitted to *Journal of Hydrometeorology*

Tian, H., Melillo, J.M., Kicklighter, D.W., McGuire, A.D. and Helfrich, J., 1999, The sensitivity of terrestrial carbon storage to historical climate variability and atmospheric CO_2 in the United States, *Tellus,* **51B**, 414–452

Viterbo, P. and Beljaars, A.C.M., 1995, An improved land surface parameterization scheme in the ECMWF model and its validation, *J. Climate,* **8**, 2,716–2,748

Xiao, X., Melillo, J.M., Kicklighter, D.W., McGuire, A.D., Prinn, R.G., Wang, C., Stone, P.H. and Sokolov, A., 1998, Transient climate change and net ecosystem production of the terrestrial biosphere. **Global Biogeochemical Cycles, 12**, 345–360

Yang, Z.-L., Dickinson, R.E. Henderson-Sellers, A., Pitman, A.J., 1995, Preliminary study of spin-up processes in land surface models with the first stage data of Project for Intercomparison of Land Surface Parameterization Schemes Phase 1(a), *J. Geophys. Res. (Atmospheres),* **100(D8)**, 16,553–16,578

Table 1. Surface energy residual (W m^{-2}) statistics from LSSs in PILPS Phase 1(b) (repeated offline intercomparisons) equilibrium climates after multi-year simulations and AMIP I/12 (coupled LSS–AGCMs) averaged over 10 years

	Mean	**Minimum**	**Maximum**
AMIP (I/12)			
Global land	3.9	−0.4	13.2
S.H. land	2.8	−1.4	11.4
N.H. land	4.4	−0.3	14.1
Offline (PILPS Phase 1(b))			
Tropical forest	0.3	−6.7	14.0
Grassland	1.5	−8.4	13.4

Figure Legends

Figure 1(a) Annually-averaged values of evapotranspiration and sensible heat fluxes (W m^{-2}) for a tropical forest environment derived from two phases of the PILPS intercomparisons: Phase 1(a), 1(b) (repeat of 1(a)) and AMIP I/12/AMIP. For each result set the cross bars show the mean values and the ranges. The point values are from PILPS Phase 1(b). AMIP values are derived from 10 models including landsurface schemes of interest. In each case, values from 9 grid points closest to those used to derive the PILPS Phase 1 forcing [37.80°N, 105°W and 6.7°S, 60°W] were averaged using a linear weighting. Individual LSSs (PILPS) and AGCMs + LSS (AMIP) are intentionally not identified.

Figure 1(b) As for (a) but for a mid-latitude grassland.

Figure 2 Zonally and annually averaged values of the surface energy residual (W m^{-2}) calculated for continental points from selected AMIP I experiments analyzed in diagnostic subproject number 12. The full 10 years of the AMIP simulations have been used. The participating AGCMs are intentionally not identified.